ELEC 474: Machine Vision Take Home Exam

Andrew Litt

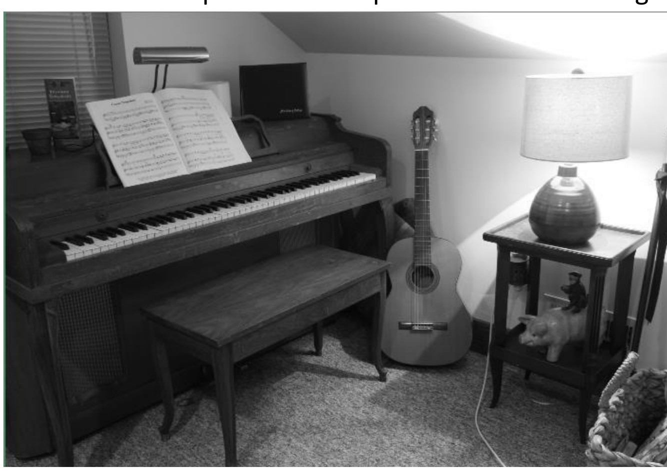
Part 1: Stereo Images to Disparity Map

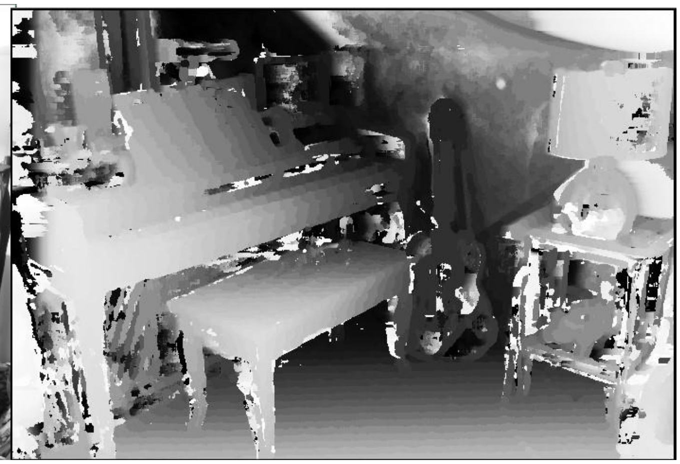
In this section, a disparity map was created through the comparison of the two provided images. Selected pixels in one frame are compared to those in a reference frame row within a 7x7 window, with displacement being determined by the comparison output with the lowest sum-of-absolute-differences (SAD). The displacement values are then stored into an image, normalized and equalized. A brighter value of the output image depicts a greater displacement between the two provided frames.

To improve computing efficiency, calculating the SAD between row pixels was limited to a displacement range given in the program, Dmax, instead of computing the SAD for each pixel in the row.

Within this section, the only OpenCV methods used were absdiff(…), normalize(…), and equalizeHistogram(…), in addition to basic conversion methods between stored datatypes.

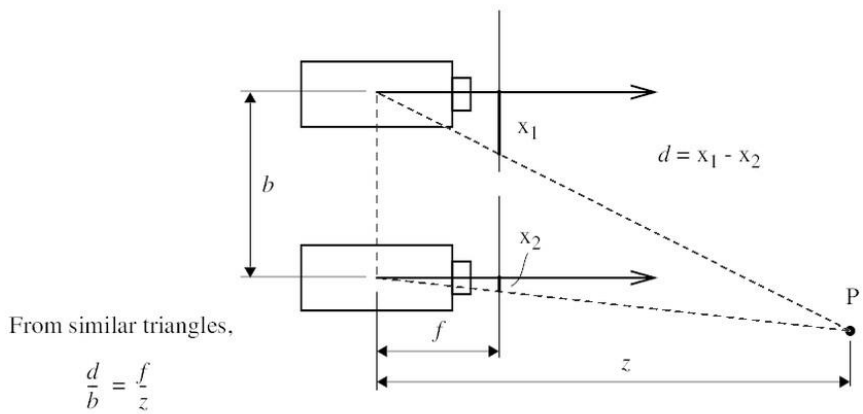
The resultant displacement maps can be viewed alongside the source image below.





Part 2: Disparity Map to 3D Point Set

To convert the created disparity map into a set of 3D points, the first step is to use disparity map data to determine real depth of an object in the scene. The figure below depicts how similar triangles can be used to determine the real depth as related to the camera focal length f, baseline b, and the determined disparity value.



From this, we can relate the real depth of an object to be:

$$\frac{displacement}{baseline} \quad \frac{focal\ length}{real\ depth} =$$

Rearranging the system, we can determine:

$$Z = \frac{focal\ length * baseline}{displacement + offset}$$

After finding Z, we can then solve for the X and Y coordinates from using the calibrated camera parameters as provided by the middlebury dataset. With Z, it is simple to rearrange the equations for X_i & Y_i, the xy coordinates of the image, to obtain X_w & X_w, the xy coordinates in 3D space.

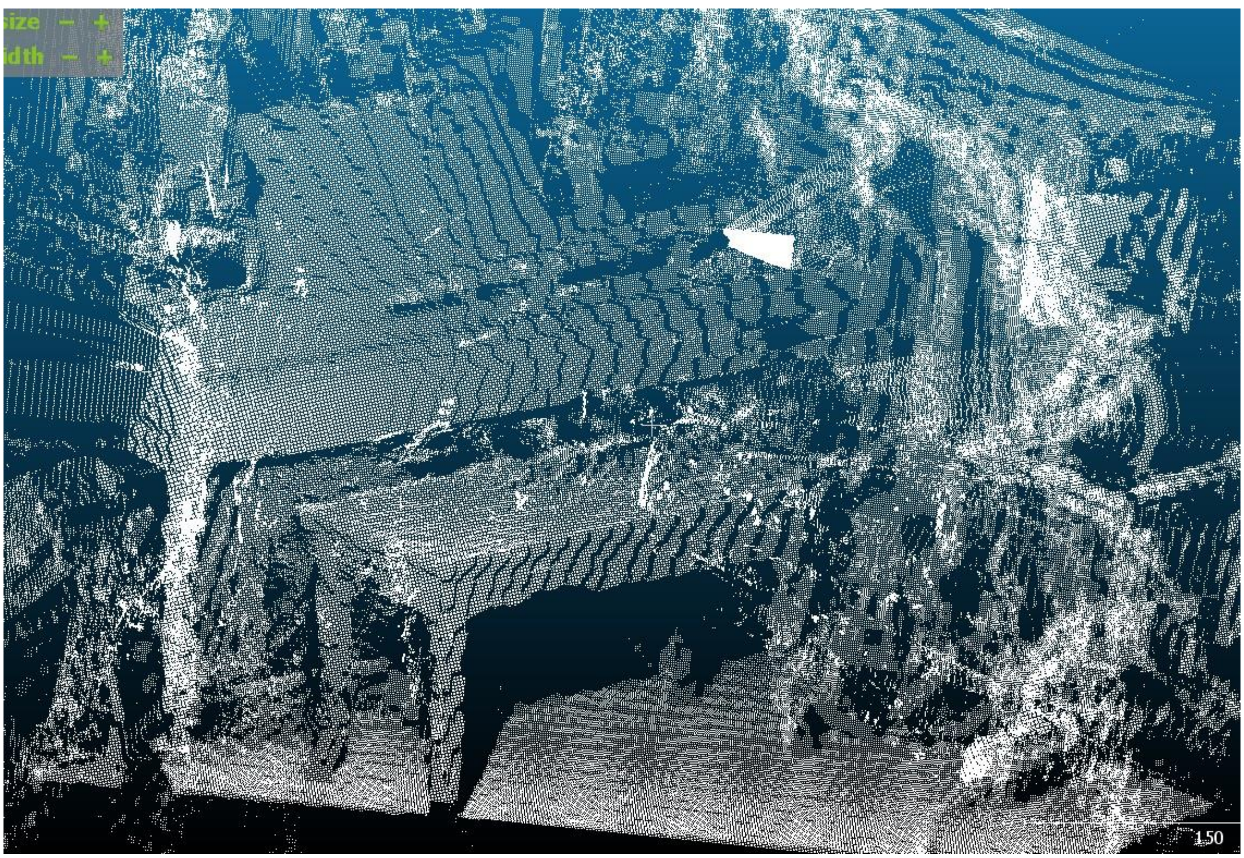
$$X_i = f * \frac{X_w}{Z} + c_x \qquad \rightarrow \qquad X_w = (X_i - c_x) * \frac{f}{Z}$$

$$Y_i = f * \frac{Y_w}{Z} + c_y \qquad \rightarrow \qquad Y_w = (Y_i - c_y) * \frac{f}{Z}$$

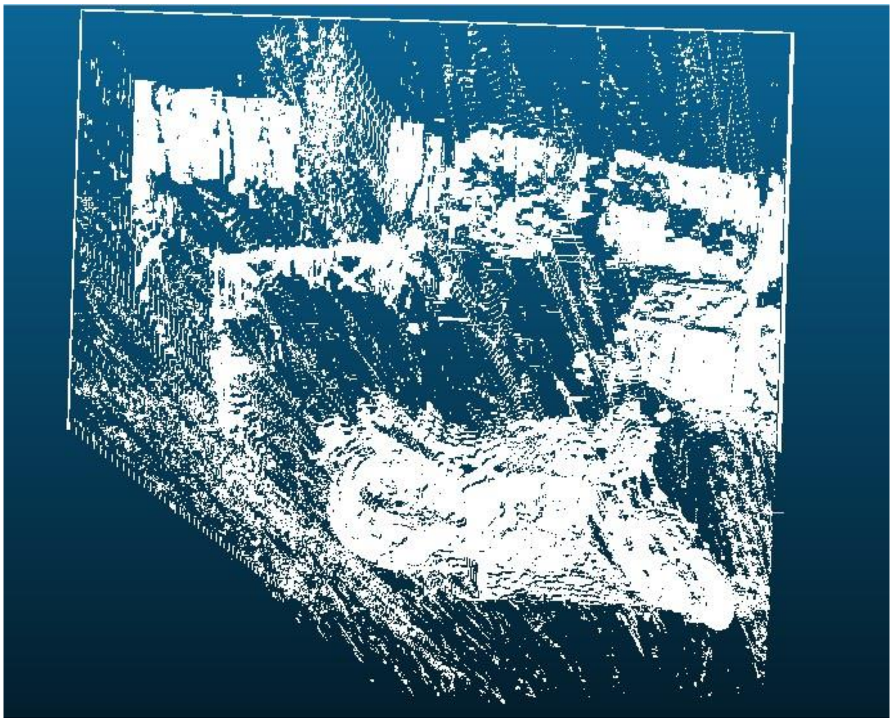
In using the provided system parameters, the conversion from x-y-disparity to points in 3D space can be calculated simply using the equations stated.

The implementation of 'no-match' values in the disparity map was never reached, which in turn causes for a large amount of noise in the finished 3D point clouds. The resultant 3D point clouds can be seen below. The resultant points in 3D space were stored in pianoOutput.txt, motorcycleOutput.txt, and umbrellaOutput.txt.

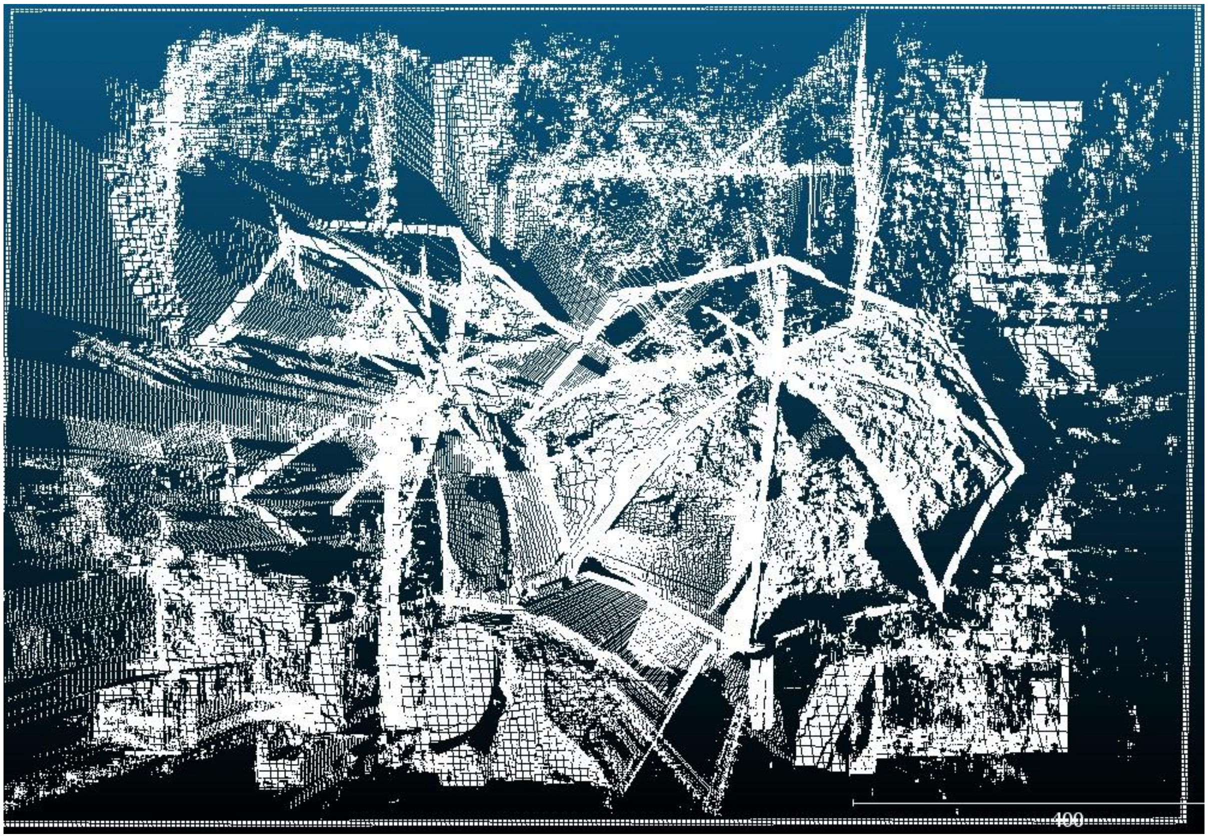
Piano Scene



Motorcycle Scene



Umbrella Scene


400